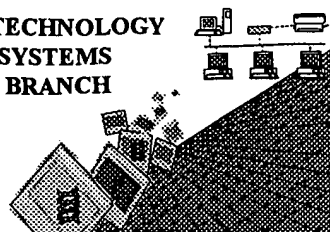


BIOTECHNOLOGY  
SYSTEMS  
BRANCH



**RAW SEQUENCE LISTING**  
**ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/020,436  
Source: OIPE  
Date Processed by STIC: 2/21/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

**<http://www.uspto.gov/web/offices/pac/checker>**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebs/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:  
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7<sup>th</sup> Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202  
Or  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER: 10/020,436

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics  
    Wrapped Aminos      The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping"
- 2      Invalid Line Length      The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3      Misaligned Amino  
    Numbering      The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4      Non-ASCII      The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 ✓ Variable Length      Sequence(s) 1-24 contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6      PatentIn 2.0  
    "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)           . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7      Skipped Sequences  
    (OLD RULES)      Sequence(s)            missing. If intentional, please insert the following lines for each skipped sequence:  
    (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
    (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
    (ii) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
    This sequence is intentionally skipped  
  
    Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8      Skipped Sequences  
    (NEW RULES)      Sequence(s)            missing. If intentional, please insert the following lines for each skipped sequence.  
    <210> sequence id number  
    <400> sequence id number  
    000
- 9      Use of n's or Xaa's  
    (NEW RULES)      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
    Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
    In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.
- 10      Invalid <213>  
    Response      Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11      Use of <220>      Sequence(s)            missing the <220> "Feature" and associated numeric identifiers and responses.  
    Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
    (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12      PatentIn 2.0  
    "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13      Misuse of n      n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.



OIPE

## RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/10/020,436

TIME: 11:29:58

Input Set : A:\39200.app

Output Set: N:\CRF3\02212002\J020436.raw

3 <110> APPLICANT: PEPTIDE THERAPEUTICS LIMITED et al  
 5 <120> TITLE OF INVENTION: A METHOD FOR MAPPING THE ACTIVE SITES BOUND BY ENZYMES  
 6 THAT COVALENTLY MODIFY SUBSTRATE MOLECULES  
 8 <130> FILE REFERENCE: 39200A/JMD/NT  
 10 <140> CURRENT APPLICATION NUMBER: US/10/020,436  
 11 <141> CURRENT FILING DATE: 2001-12-18  
 13 <150> PRIOR APPLICATION NUMBER: GB 9722818.3  
 14 <151> PRIOR FILING DATE: 1997-10-30  
 16 <160> NUMBER OF SEQ ID NOS: 21  
 18 <170> SOFTWARE: PatentIn Ver. 2.1  
 20 <210> SEQ ID NO: 1  
 21 <211> LENGTH: 3  
 22 <212> TYPE: PRT  
 23 <213> ORGANISM: Artificial Sequence  
 25 <220> FEATURE:  
 26 <223> OTHER INFORMATION: Description of Artificial Sequence: formula for a  
 27 library comprising natural and/or unnatural amino  
 28 acid residues or peptidomimetics  
 30 <220> FEATURE:  
 31 <221> NAME/KEY: UNSURE  
 32 <222> LOCATION: (1)  
 33 <223> OTHER INFORMATION: Xaa at position 1 is any natural or unnatural  
 34 amino acid residue or peptidomimetic and is  
 35 repeated x times  
 37 <220> FEATURE:  
 38 <221> NAME/KEY: UNSURE  
 39 <222> LOCATION: (2)  
 40 <223> OTHER INFORMATION: Xaa at position 2 is a non-degenerate modifiable  
 41 natural or unnatural amino acid residue or  
 42 peptidomimetic  
 44 <220> FEATURE:  
 45 <221> NAME/KEY: UNSURE  
 46 <222> LOCATION: (3)  
 47 <223> OTHER INFORMATION: Xaa at position 3 is any natural or unnatural  
 48 amino acid residue or peptidomimetic and is  
 49 repeated y times.  
 51 <220> FEATURE:  
 52 <221> NAME/KEY: UNSURE  
 53 <222> LOCATION: (1)..(3)  
 54 <223> OTHER INFORMATION: x and y are each independently 0 or an integer; (x  
 55 + y) = (n - 1); and n = an integer from 3 to 8,  
 56 preferably 5  
 58 <400> SEQUENCE: 1

Does Not Comply  
Corrected Diskette Needed

pp 1-3

invalid - see item 5 on  
Err Summary sheet

same error  
↓

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/020,436

DATE: 02/21/2002  
TIME: 11:29:58

Input Set : A:\39200.app  
Output Set : N:\CRF3\02212002\J020436.raw

W--> 59 Xaa Xaa Xaa  
60 1  
63 <210> SEQ ID NO: 2  
64 <211> LENGTH: 3  
65 <212> TYPE: PRT  
66 <213> ORGANISM: Artificial Sequence  
68 <220> FEATURE:  
69 <223> OTHER INFORMATION: Description of Artificial Sequence: formula for  
70 library used to identify protein kinase inhibitor  
71 molecules  
73 <220> FEATURE:  
74 <221> NAME/KEY: UNSURE  
75 <222> LOCATION: (1)  
76 <223> OTHER INFORMATION: Xaa at position 1 is any natural or unnatural  
77 amino acid residue or peptidomimetic and is  
78 repeated x times *item 5*  
80 <220> FEATURE:  
81 <221> NAME/KEY: UNSURE  
82 <222> LOCATION: (3)  
83 <223> OTHER INFORMATION: Xaa at position 3 is any natural or unnatural  
84 amino acid residue or peptidomimetic and is  
85 repeated y times *item 5*  
87 <220> FEATURE:  
88 <221> NAME/KEY: UNSURE  
89 <222> LOCATION: (1)..(3)  
90 <223> OTHER INFORMATION: x and y are each independently 0 or an integer; (x  
91 + y) = (n-1); and n = an integer from 3 to 8,  
92 preferably 5  
94 <400> SEQUENCE: 2  
W--> 95 Xaa Tyr Xaa  
96 1  
99 <210> SEQ ID NO: 3  
100 <211> LENGTH: 8  
101 <212> TYPE: PRT  
102 <213> ORGANISM: Artificial Sequence  
104 <220> FEATURE:  
105 <223> OTHER INFORMATION: Description of Artificial Sequence: consensus  
106 peptide substrate for ZAP-70  
108 <220> FEATURE:  
109 <221> NAME/KEY: UNSURE  
110 <222> LOCATION: (1)  
111 <223> OTHER INFORMATION: Biotin-epsilon-amino hexanoic acid is linked to the  
112 aspartic acid residue at position 1  
114 <220> FEATURE:  
115 <221> NAME/KEY: UNSURE  
116 <222> LOCATION: (8)  
117 <223> OTHER INFORMATION: Leucine at position 8 is Norleucine  
119 <400> SEQUENCE: 3  
120 Asp Glu Glu Asp Tyr Phe Glu Leu

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/020,436

DATE: 02/21/2002

TIME: 11:29:58

Input Set : A:\39200.app

Output Set: N:\CRF3\02212002\J020436.raw

121 1 5  
 124 <210> SEQ ID NO: 4  
 125 <211> LENGTH: 3  
 126 <212> TYPE: PRT  
 127 <213> ORGANISM: Artificial Sequence  
 129 <220> FEATURE:  
 130 <223> OTHER INFORMATION: Description of Artificial Sequence: general  
 131 formula for peptide library containing substrates  
 132 for protein serine or serine/threonine kinase  
 133 enzyme  
 135 <220> FEATURE:  
 136 <221> NAME/KEY: UNSURE  
 137 <222> LOCATION: (1)  
 138 <223> OTHER INFORMATION: Xaa at position 1 is any natural or unnatural  
 139 amino acid residue or peptidomimetic and is  
 140 repeated x times *item 5*  
 142 <220> FEATURE:  
 143 <221> NAME/KEY: UNSURE  
 144 <222> LOCATION: (3)  
 145 <223> OTHER INFORMATION: Xaa at position 3 is any natural or unnatural  
 146 amino acid residue or peptidomimetic and is  
 147 repeated y times  
 149 <220> FEATURE:  
 150 <221> NAME/KEY: UNSURE *item 5*  
 151 <222> LOCATION: (1)..(3)  
 152 <223> OTHER INFORMATION: x and y are each independently 0 or an integer; (x  
 153 + y) = (n - 1); and n = an integer from 3 to 8,  
 154 preferably 5  
 156 <400> SEQUENCE: 4  
 W--> 157 Xaa Ser Xaa  
 158 1  
 161 <210> SEQ ID NO: 5  
 162 <211> LENGTH: 36  
 163 <212> TYPE: DNA  
 164 <213> ORGANISM: Artificial Sequence  
 166 <220> FEATURE:  
 167 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR primer  
 168 used to amplify the coding sequence for human  
 169 ZAP-70 amino acids 306-615 from Jurkat T cell cDNA  
 171 <400> SEQUENCE: 5  
 172 ccgggatccg ccatgcccat ggacacgagc gtgtat 36  
 174 <210> SEQ ID NO: 6  
 175 <211> LENGTH: 57  
 176 <212> TYPE: DNA  
 177 <213> ORGANISM: Artificial Sequence  
 179 <220> FEATURE:  
 180 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR primer  
 181 used to amplify the coding sequence for human  
 182 ZAP-70 amino acids 306-615 from Jurkat T cell cDNA

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/020,436

DATE: 02/21/2002

TIME: 11:29:58

Input Set : A:\39200.app

Output Set: N:\CRF3\02212002\J020436.raw

```

184 <400> SEQUENCE: 6
185 gggggatcct cagtgggtggt ggtgggtggtg ggcacaggca gcctcagcct tctgtgt      57
187 <210> SEQ ID NO: 7
188 <211> LENGTH: 5
189 <212> TYPE: PRT
190 <213> ORGANISM: Artificial Sequence
192 <220> FEATURE:
193 <223> OTHER INFORMATION: Description of Artificial Sequence: sequence of
194     phosphorylated motif identified by screen of
195     library peptides
197 <400> SEQUENCE: 7
198 Asp Glu Glu Asp Tyr
199   1           5
202 <210> SEQ ID NO: 8
203 <211> LENGTH: 5
204 <212> TYPE: PRT
205 <213> ORGANISM: Artificial Sequence
207 <220> FEATURE:
208 <223> OTHER INFORMATION: Description of Artificial Sequence: sequence of
209     phosphorylated motif identified by screen of
210     library peptides
212 <400> SEQUENCE: 8
213 Asp Glu Glu Tyr Phe
214   1           5
217 <210> SEQ ID NO: 9
218 <211> LENGTH: 5
219 <212> TYPE: PRT
220 <213> ORGANISM: Artificial Sequence
222 <220> FEATURE:
223 <223> OTHER INFORMATION: Description of Artificial Sequence: sequence of
224     phosphorylated peptide motif identified by screen
225     of library peptides
227 <400> SEQUENCE: 9
228 Asp Glu Tyr Glu Phe
229   1           5
232 <210> SEQ ID NO: 10
233 <211> LENGTH: 5
234 <212> TYPE: PRT
235 <213> ORGANISM: Artificial Sequence
237 <220> FEATURE:
238 <221> NAME/KEY: UNSURE
239 <222> LOCATION: (5)
240 <223> OTHER INFORMATION: Leucine at position 5 is Norleucine
242 <220> FEATURE:
243 <223> OTHER INFORMATION: Description of Artificial Sequence: sequence of
244     phosphorylated motif identified by screen of
245     library peptides
247 <400> SEQUENCE: 10
248 Asp Tyr Phe Glu Leu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/020,436

DATE: 02/21/2002

TIME: 11:29:58

Input Set : A:\39200.app

Output Set: N:\CRF3\02212002\J020436.raw

249 1 5  
252 <210> SEQ ID NO: 11  
253 <211> LENGTH: 5  
254 <212> TYPE: PRT  
255 <213> ORGANISM: Artificial Sequence  
257 <220> FEATURE:  
258 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide  
259 identified as a preferred substrate for Syk  
261 <400> SEQUENCE: 11  
262 Asp Glu Glu Asp Tyr  
263 1 5  
266 <210> SEQ ID NO: 12  
267 <211> LENGTH: 5  
268 <212> TYPE: PRT  
269 <213> ORGANISM: Artificial Sequence  
271 <220> FEATURE:  
272 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide  
273 identified as a preferred substrate for Syk  
275 <400> SEQUENCE: 12  
276 Asp Glu Glu Tyr Asp  
277 1 5  
280 <210> SEQ ID NO: 13  
281 <211> LENGTH: 5  
282 <212> TYPE: PRT  
283 <213> ORGANISM: Artificial Sequence  
285 <220> FEATURE:  
286 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide  
287 identified as a preferred substrate for Syk  
289 <400> SEQUENCE: 13  
290 Asp Glu Tyr Glu Asp  
291 1 5  
294 <210> SEQ ID NO: 14  
295 <211> LENGTH: 5  
296 <212> TYPE: PRT  
297 <213> ORGANISM: Artificial Sequence  
299 <220> FEATURE:  
300 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide  
301 identified as a preferred substrate for Syk  
303 <400> SEQUENCE: 14  
304 Asp Tyr Glu Glu Val  
305 1 5  
308 <210> SEQ ID NO: 15  
309 <211> LENGTH: 5  
310 <212> TYPE: PRT  
311 <213> ORGANISM: Artificial Sequence  
313 <220> FEATURE:  
314 <221> NAME/KEY: UNSURE  
315 <222> LOCATION: (5)  
316 <223> OTHER INFORMATION: Leucine at position 5 is Norleucine

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/020,436

DATE: 02/21/2002

TIME: 11:29:59

Input Set : A:\39200.app

Output Set: N:\CRF3\02212002\J020436.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:59 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:95 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:157 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4